LARGE DEVIATIONS FOR MULTI-VALUED STOCHASTIC DIFFERENTIAL EQUATIONS*

JIAGANG REN[1], SIYAN XU[2], XICHENG ZHANG[2,3]

[1] *School of Mathematics and Computational Science, Sun Yat-Sen University, Guangzhou, Guangdong 510275, P.R.China*
[2]*School of Mathematics and statistic, Huazhong University of Science and Technology, Wuhan, Hubei 430074, P.R.China*
[3]*School of Mathematics and Statistics, The University of New South Wales, Sydney, 2052, Australia*
Emails: J. Ren: renjg@mail.sysu.edu.cn
S. Xu: xsy_00@hotmail.com
X.Zhang: xichengzhang@gmail.com

ABSTRACT. We prove a large deviation principle of Freidlin-Wentzell's type for the multi-valued stochastic differential equations with monotone drifts, which in particular contains a class of SDEs with reflection in a convex domain.

1. INTRODUCTION

Consider the following multivalued stochastic differential equation (MSDE in short):

$$\begin{cases} \mathrm{d}X(t) \in b(X(t))\mathrm{d}t + \sigma(X(t))\mathrm{d}W(t) - A(X(t))\mathrm{d}t, \\ X(0) = x \in \overline{D(A)}, \end{cases} \tag{1}$$

where A is a multivalued maximal monotone operator, which will be described below, $W(t) = \{W^k(t), t \geqslant 0, k \in \mathbb{N}\}$ is a sequence of independent standard Brownian motions on a filtered probability space $(\Omega, \mathcal{F}, P; (\mathcal{F}_t)_{t \geqslant 0})$, $b : \mathbb{R}^m \to \mathbb{R}^m$ and $\sigma : \mathbb{R}^m \to \mathbb{R}^m \times l^2$ are two continuous functions, l^2 stands for the Hilbert space of square summable sequences of real numbers.

This type of MSDE was first studied by Cépa in [5, 6]. He proved that if b and σ are Lipschitz continuous, then there exists a unique pair of processes $(X(t), K(t))$ such that

$$X(t) = x + \int_0^t b(X(s))\mathrm{d}s + \int_0^t \sigma(X(s))\mathrm{d}W(s) - K(t),$$

where $K(t)$ is a process of finite variation (see Definition 2.3 below for more details). Recently, Zhang [14] extended Cépa's result to the infinite dimensional case, and relaxed the Lipschitz assumption on b to the monotone case. It should be noted that when A is the subdifferential of the indicator function of a convex subset of $\mathbb{R}^m$, the above MSDE is the same as the usual SDE with reflecting boundary in a convex domain (cf. [1, 9]). Moreover, since the subdifferential of any lower semicontinuous convex function is a maximal monotone operator, Cépa's result can also be used to deal with the SDE with discontinuous coefficients. It is well known that there are many literatures to investigate the SDEs with reflecting boundary since the solutions of a large class of PDEs with

Neumann boundary and mixed boundary conditions can be represented by the solution of such SDEs (cf. [1]).

We now consider the following small perturbation of Eq.(1):

$$\begin{cases} \mathrm{d}X^\epsilon(t) \in b(X^\epsilon(t))\mathrm{d}t + \sqrt{\epsilon}\sigma(X^\epsilon(t))\mathrm{d}W(t) - A(X^\epsilon(t))\mathrm{d}t, \\ X^\epsilon(0) = x \in \overline{D(A)}, \quad \epsilon \in (0,1]. \end{cases} \tag{2}$$

The solution of this equation is denoted by $(X^\epsilon(t,x), K^\epsilon(t,x))$. We want to establish the large deviation principle of the law of $X^\epsilon(t,x)$ in the space $\mathbb{S} := C([0,T] \times \overline{D(A)}; \overline{D(A)})$, namely, the asymptotic estimates of probabilities $P(X^\epsilon \in \Gamma)$, where $\Gamma \in \mathcal{B}(\mathbb{S})$.

In [1], Anderson and Orey considered the same small random perturbation for the dynamical system with reflecting boundary in smooth domain, and obtained the Freidlin-Wentzell's large deviation estimates in $C([0,T]; \overline{D(A)})$. They assumed that the coefficients are bounded and Lipschitz continuous, and the diffusion coefficient is non-degenerate. Using the contraction principle, Cépa [5] only considered the large deviation principle of one dimensional case based on an explicit construction of the solution (cf.[13]). The multi-dimensional case is still open. Compared with the usual SDE, i.e., $A = 0$, most of the difficulties come from the presence of the process of finite variation, $K(t)$. One only knows that $t \mapsto K(t)$ is continuous, and could not prove any further regularity such as Hölder continuity. Therefore, the classical method of time discretization is almost inapplicable (cf. [7]).

Our method is based on the recently well developed weak convergence approach due to Dupuis and Ellis [8] (see also [2, 3]). This method has been proved to be very effective for various systems (cf. [11, 15, 4, 16, 10, 12, etc.]). In the situation considered in the present paper, however, since we cannot prove the following uniform estimate as in [11]: for any $p \geqslant 2$ and $s, t \in [0,T]$, $x, y \in \overline{D(A)}$

$$\sup_{\epsilon \in (0,1)} \mathbb{E}|X^\epsilon(t,x) - X^\epsilon(s,y)|^{2p} \leqslant C(|t-s|^p + |x-y|^{2p}),$$

we cannot obtain the tightness of the laws of $X^\epsilon(t,x)$ in $\mathbb{S}$. Some technical difficulties for verifying the conditions $(\mathbf{LD})_1$ and $(\mathbf{LD})_2$ below need to be overcome.

In Section 2, we recall some well known facts about the MSDE and a criterion for Laplace principle. In Section 3, we present our main result and give a detailed proof. Throughout the paper, C with or without indexes will denote different constants (depending on the indexes) whose values are not important.

2. Preliminaries

We first give some notions and notations about multivalued operators. Let $2^{\mathbb{R}^m}$ be the set of all subsets of $\mathbb{R}^m$. A map $A : \mathbb{R}^m \to 2^{\mathbb{R}^m}$ is called a multivalued operator. Given such a multivalued operator A, define:

$$\begin{aligned} D(A) &:= \{x \in \mathbb{R}^m : A(x) \neq \emptyset\}, \\ \mathrm{Im}(A) &:= \cup_{x \in D(A)} A(x), \\ \mathrm{Gr}(A) &:= \{(x,y) \in \mathbb{R}^{2m} : x \in \mathbb{R}^m, y \in A(x)\}. \end{aligned}$$

We recall the following definitions.

Definition 2.1. *(1) A multivalued operator A is called monotone if*

$$\langle y_1 - y_2, x_1 - x_2 \rangle_{\mathbb{R}^m} \geqslant 0, \quad \forall (x_1, y_1), (x_2, y_2) \in \mathrm{Gr}(A).$$

(2) A monotone operator A is called maximal monotone if for each $(x, y) \in \mathrm{Gr}(A)$,

$$\langle y - y', x - x' \rangle_{\mathbb{R}^m} \geqslant 0, \quad \forall (x', y') \in \mathrm{Gr}(A).$$

Examples 2.2. *Suppose that $\mathcal{O}$ is a closed convex subset of $\mathbb{R}^m$, and $I_{\mathcal{O}}$ is the indicator function of $\mathcal{O}$, i.e,*

$$I_{\mathcal{O}}(x) = \begin{cases} 0, & \text{if } x \in \mathcal{O}, \\ +\infty, & \text{if } x \notin \mathcal{O}. \end{cases}$$

The subdifferential of $I_{\mathcal{O}}$ is given by

$$\begin{aligned} \partial I_{\mathcal{O}}(x) &= \{ y \in \mathbb{R}^m : \langle y, x - z \rangle_{\mathbb{R}^m} \geqslant 0, \forall z \in \mathcal{O} \} \\ &= \begin{cases} \emptyset, & \text{if } x \notin \mathcal{O}, \\ \{0\}, & \text{if } x \in \mathrm{Int}(\mathcal{O}), \\ \Lambda_x, & \text{if } x \in \partial \mathcal{O}, \end{cases} \end{aligned}$$

where $\mathrm{Int}(\mathcal{O})$ is the interior of $\mathcal{O}$ and Λ_x is the exterior normal cone at x. One can check that $\partial I_{\mathcal{O}}$ is a multivalued maximal monotone operator in the sense of Definition 2.1.

We now give the precise definition of the solution to Eq.(1).

Definition 2.3. *A pair of continuous and $(\mathcal{F}_t)$-adapted processes (X, K) is called a solution of Eq.(1) if*

(i) $X(0) = x$, and for all $t \geqslant 0$, $X(t) \in \overline{D(A)}$ a.s.;
(ii) $K(0) = 0$ a.s. and K is of finite variation;
(iii) $\mathrm{d}X(t) = b(X(t))\mathrm{d}t + \sigma(X(t))\mathrm{d}W(t) - \mathrm{d}K(t), \ 0 \leqslant t < \infty,$ a.s.;
(iv) for any continuous and $(\mathcal{F}_t)$-adapted processes (α, β) with

$$(\alpha(t), \beta(t)) \in \mathrm{Gr}(A), \quad \forall t \in [0, +\infty),$$

the measure

$$\langle X(t) - \alpha(t), \mathrm{d}K(t) - \beta(t)\mathrm{d}t \rangle \geqslant 0 \quad a.s..$$

We now recall an abstract criterion for Laplace principle, which is equivalent to the large deviation principle (cf. [3, 4, 15]). It is well known that there exists a Hilbert space so that $l^2 \subset \mathbb{U}$ is Hilbert-Schmidt with embedding operator J and $\{W^k(t), k \in \mathbb{N}\}$ is a Brownian motion with values in $\mathbb{U}$, whose covariance operator is given by $Q = J \circ J^*$. For example, one can take $\mathbb{U}$ as the completion of l^2 with respect to the norm generated by the scalar product

$$\langle h, h' \rangle_{\mathbb{U}} := \left(\sum_{k=1}^{\infty} \frac{h_k h'_k}{k^2} \right)^{\frac{1}{2}}, \quad h, h' \in l^2.$$

For a Polish space $\mathbb{B}$, we denote by $\mathcal{B}(\mathbb{B})$ its Borel σ-field, and by $\mathbb{C}_T(\mathbb{B})$ the continuous function space from $[0, T]$ to $\mathbb{B}$, which is endowed with the uniform distance so that $\mathbb{C}_T(\mathbb{B})$ is still a Polish space. Define

$$\ell_T^2 := \left\{ h = \int_0^{\cdot} \dot{h}(s)\mathrm{d}s : \ \dot{h} \in L^2(0, T; l^2) \right\} \tag{3}$$

with the norm

$$\|h\|_{\ell_T^2} := \left(\int_0^T \|\dot{h}(s)\|_{l^2}^2 \mathrm{d}s \right)^{1/2},$$

where the dot denotes the generalized derivative. Let μ be the law of the Brownian motion W in $\mathbb{C}_T(\mathbb{U})$. Then

$$(\ell_T^2, \mathbb{C}_T(\mathbb{U}), \mu)$$

forms an abstract Wiener space.

For $T, N > 0$, set

$$\mathcal{D}_N := \{h \in \ell_T^2 : \|h\|_{\ell_T^2} \leqslant N\}$$

and

$$\mathcal{A}_N^T := \left\{ \begin{array}{l} h : [0, T] \to l^2 \text{ is a continuous and } (\mathcal{F}_t)\text{-adapted} \\ \text{process, and for almost all } \omega, \quad h(\cdot, \omega) \in \mathcal{D}_N \end{array} \right\}. \tag{4}$$

We equip $\mathcal{D}_N$ with the weak convergence topology in ℓ_T^2. Then

$$\mathcal{D}_N \text{ is metrizable as a compact Polish space.} \tag{5}$$

Let $\mathbb{S}$ be a Polish space. A function $I : \mathbb{S} \to [0, \infty]$ is given.

Definition 2.4. *The function I is called a rate function if for every $a < \infty$, the set $\{f \in \mathbb{S} : I(f) \leqslant a\}$ is compact in $\mathbb{S}$.*

Let $\{Z^\epsilon : \mathbb{C}_T(\mathbb{U}) \to \mathbb{S}, \epsilon \in (0, 1)\}$ be a family of measurable mappings. Assume that there is a measurable map $Z_0 : \ell_T^2 \mapsto \mathbb{S}$ such that

(LD)₁ For any $N > 0$, if a family $\{h_\epsilon, \epsilon \in (0, 1)\} \subset \mathcal{A}_N^T$ (as random variables in $\mathcal{D}_N$) converges in distribution to $h \in \mathcal{A}_N^T$, then for some subsequence ϵ_k, $Z^{\epsilon_k}\left(\cdot + \frac{h_{\epsilon_k}(\cdot)}{\sqrt{\epsilon_k}}\right)$ converges in distribution to $Z_0(h)$ in $\mathbb{S}$.

(LD)₂ For any $N > 0$, if $\{h_n, n \in \mathbb{N}\} \subset \mathcal{D}_N$ weakly converges to $h \in \ell_T^2$, then for some subsequence h_{n_k}, $Z_0(h_{n_k})$ converges to $Z_0(h)$ in $\mathbb{S}$.

For each $f \in \mathbb{S}$, define

$$I(f) := \frac{1}{2} \inf_{\{h \in \ell_T^2 : \ f = Z_0(h)\}} \|h\|_{\ell_T^2}^2, \tag{6}$$

where $\inf \emptyset = \infty$ by convention. Then under **(LD)₂**, $I(f)$ is a rate function.

We recall the following result due to [3] (see also [17, Theorem 4.4]).

Theorem 2.5. *Under **(LD)₁** and **(LD)₂**, $\{Z^\epsilon, \epsilon \in (0, 1)\}$ satisfies the Laplace principle with the rate function $I(f)$ given by (6). More precisely, for each real bounded continuous function g on $\mathbb{S}$:*

$$\lim_{\epsilon \to 0} \epsilon \log \mathbb{E}^\mu \left(\exp \left[-\frac{g(Z^\epsilon)}{\epsilon} \right] \right) = -\inf_{f \in \mathbb{S}} \{g(f) + I(f)\}. \tag{7}$$

In particular, the family of $\{Z^\epsilon, \epsilon \in (0, 1)\}$ satisfies the large deviation principle in $(\mathbb{S}, \mathcal{B}(\mathbb{S}))$ with the rate function $I(f)$. More precisely, let ν_ϵ be the law of Z^ϵ in $(\mathbb{S}, \mathcal{B}(\mathbb{S}))$, then for any $B \in \mathcal{B}(\mathbb{S})$

$$-\inf_{f \in B^o} I(f) \leqslant \liminf_{\epsilon \to 0} \epsilon \log \nu_\epsilon(B) \leqslant \limsup_{\epsilon \to 0} \epsilon \log \nu_\epsilon(B) \leqslant -\inf_{f \in \bar{B}} I(f),$$

where the closure and the interior are taken in $\mathbb{S}$, and $I(f)$ is defined by (6).

3. Main Result and Proof

We assume that

(H1) A is a maximal monotone operator with non-empty interior, i.e., $\text{Int}(D(A)) \neq \emptyset$;

4

(H2) σ and b are continuous functions and satisfy that for some $C_\sigma, C_b > 0$ and all $x, y \in \mathbb{R}^m$

$$\|\sigma(x) - \sigma(y)\|_{L_2(l^2;\mathbb{R}^m)} \leqslant C_\sigma |x - y|,$$
$$\langle x - y, b(x) - b(y)\rangle_{\mathbb{R}^m} \leqslant C_b |x - y|^2,$$

where $L_2(l^2; \mathbb{R}^m)$ denotes the Hilbert-Schmidt space and $|\cdot|$ denotes the norm in $\mathbb{R}^m$, and for some $C_b' > 0$ and $n \in \mathbb{N}$

$$|b(x)| \leqslant C_b'(1 + |x|^n).$$

It is well known that under **(H1)** and **(H2)**, there exists a unique solution (X^ϵ, K^ϵ) to Eq.(2) in the sense of Definition 2.3 (cf. [14]). Our main result is stated as follows:

Theorem 3.1. *Assume that* **(H1)** *and* **(H2)** *hold. Then the family of $\{X^\epsilon(t, x),\ \epsilon \in (0, 1)\}$ satisfies the large deviation principle in $\mathbb{S} := C([0, T] \times \overline{D(A)}; \overline{D(A)})$ with the rate function given by*

$$I(f) := \frac{1}{2} \inf_{\{h \in \ell_T^2:\ f = X^h\}} \|h\|_{\ell_T^2}^2, \tag{8}$$

where $X^h(t, x)$ solves the following equation:

$$\mathrm{d}X^h(t) \in b(X^h(t))\mathrm{d}t + \sigma(X^h(t))\dot{h}(t)\mathrm{d}t - A(X^h(t))\mathrm{d}t, \quad X^h(0) = x.$$

For proving this result, by Theorem 2.5, the main task is to verify **(LD)$_1$** and **(LD)$_2$** with

$$\mathbb{S} := C([0, T] \times \overline{D(A)}; \overline{D(A)}), \quad Z^\epsilon = X^\epsilon, \quad Z_0(h) = X^h.$$

This will be done in Lemmas 3.7 and 3.8 below.

Let $h_\epsilon \in \mathcal{A}_N^T$ converge almost surely to $h \in \mathcal{A}_N^T$ as random variables in ℓ_T^2, and $(X^{\epsilon,h_\epsilon}, K^{\epsilon,h_\epsilon})$ solve the following control equation:

$$\begin{aligned} X^{\epsilon,h_\epsilon}(t) &= x + \int_0^t b(X^{\epsilon,h_\epsilon}(s))\mathrm{d}s + \int_0^t \sigma(X^{\epsilon,h_\epsilon}(s))\dot{h}_\epsilon(s)\mathrm{d}s \\ &\quad + \sqrt{\epsilon} \int_0^t \sigma(X^{\epsilon,h_\epsilon}(s))\mathrm{d}W(s) - K^{\epsilon,h_\epsilon}(t), \end{aligned} \tag{9}$$

which can be solved by Girsanov's theorem, and (X^h, K^h) solve the following deterministic equation:

$$X^h(t) = x + \int_0^t b(X^h(s))\mathrm{d}s + \int_0^t \sigma(X^h(s))\dot{h}(s)\mathrm{d}s - K^h(t). \tag{10}$$

Let $|K|_t^s$ denote the total variation of K on $[s, t]$. We recall the following result due to Cépa [6] (see also [14, Propositions 3.3 and 3.4]).

Proposition 3.2. *Under* **(H1)**, *there exist $a \in \mathbb{R}^m, \gamma > 0, \mu \geqslant 0$ such that for any pair of (X, K) with the property (iv) of Definition 2.3 and all $0 \leqslant s < t \leqslant T$*

$$\int_s^t \langle X(r) - a, \mathrm{d}K(r)\rangle_{\mathbb{R}^m} \geqslant \gamma |K|_t^s - \mu \int_s^t |X(r) - a|\mathrm{d}r - \gamma\mu(t - s). \tag{11}$$

Moreover, for any pairs of (X, K) and $(\tilde{X}, \tilde{K})$ with the property (iv) of Definition 2.3

$$\langle X(t) - \tilde{X}(t), \mathrm{d}K(t) - \mathrm{d}\tilde{K}(t)\rangle_{\mathbb{R}^m} \geqslant 0. \tag{12}$$

Using this property, we first prove the following uniform estimates.

Lemma 3.3. *For any $p \geqslant 1$, there exists $C_{p,T,N} > 0$ such that for any $\epsilon \in (0,1)$ and $x, y \in \overline{D(A)}$*

$$\mathbb{E}\left(\sup_{t \in [0,T]} |X^{\epsilon,h_\epsilon}(t,x) - X^{\epsilon,h_\epsilon}(t,y)|^{2p}\right) \leqslant C_{p,T,N}|x-y|^{2p}. \tag{13}$$

Proof. Set

$$Z_\epsilon(t) := X^{\epsilon,h_\epsilon}(t,x) - X^{\epsilon,h_\epsilon}(t,y)$$

and

$$\Lambda(s) := \sigma(X^{\epsilon,h_\epsilon}(s,x)) - \sigma(X^{\epsilon,h_\epsilon}(s,y)).$$

By Itô's formula, **(H2)** and (12), we have for any $p \geqslant 1$

$$
\begin{aligned}
|Z_\epsilon(t)|^{2p} &= |Z_\epsilon(0)|^{2p} + 2p\int_0^t |Z_\epsilon(s)|^{2p-2}\langle Z_\epsilon(s), b(X^{\epsilon,h_\epsilon}(s,x)) - b(X^{\epsilon,h_\epsilon}(s,y))\rangle_{\mathbb{R}^m}\mathrm{d}s \\
&\quad + 2p\int_0^t |Z_\epsilon(s)|^{2p-2}\langle Z_\epsilon(s), \Lambda(s)\dot{h}_\epsilon(s)\rangle_{\mathbb{R}^m}\mathrm{d}s \\
&\quad + 2p\sqrt{\epsilon}\int_0^t |Z_\epsilon(s)|^{2p-2}\langle Z_\epsilon(s), \Lambda(s)\mathrm{d}W(s)\rangle_{\mathbb{R}^m} \\
&\quad - 2p\int_0^t |Z_\epsilon(s)|^{2p-2}\langle Z_\epsilon(s), \mathrm{d}K^{\epsilon,h_\epsilon}(s,x) - \mathrm{d}K^{\epsilon,h_\epsilon}(s,y)\rangle_{\mathbb{R}^m} \\
&\quad - p\int_0^t |Z_\epsilon(s)|^{2p-2}(\|\Lambda(s)\|^2 + 2(p-1)\langle Z_\epsilon(s), \Lambda(s)\Lambda^*(s)Z_\epsilon(s)\rangle_{\mathbb{R}^m}/|Z_\epsilon(s)|^2)\mathrm{d}s \\
&\leqslant |Z_\epsilon(0)|^{2p} + C\int_0^t |Z_\epsilon(s)|^{2p}\mathrm{d}s + 2p\int_0^t |Z_\epsilon(s)|^{2p} \cdot \|\dot{h}_\epsilon(s)\|_{l^2}\mathrm{d}s \\
&\quad + 2p\sqrt{\epsilon}\int_0^t |Z_\epsilon(s)|^{2p-2}\langle Z_\epsilon(s), \Lambda(s)\mathrm{d}W(s)\rangle_{\mathbb{R}^m}.
\end{aligned}
$$

Set

$$g(t) := \mathbb{E}\left(\sup_{s \in [0,t]} |Z_\epsilon(s)|^{2p}\right).$$

By BDG's inequality and Young's inequality, we have for any $\delta > 0$

$$
\begin{aligned}
\mathbb{E}\left|\sup_{t' \in [0,t]}\int_0^{t'} |Z_\epsilon(s)|^{2p-2}\langle Z_\epsilon(s), \Lambda(s)\mathrm{d}W(s)\rangle_{\mathbb{R}^m}\right| \\
&\leqslant C\mathbb{E}\left(\int_0^t |Z_\epsilon(s)|^{4p-4}\|\Lambda(s)^*Z_\epsilon(s)\|_{l^2}^2\mathrm{d}s\right)^{1/2} \\
&\leqslant C\mathbb{E}\left(\sup_{s \in [0,t]} |Z_\epsilon(s)|^{2p}\int_0^t |Z_\epsilon(s)|^{2p}\mathrm{d}s\right)^{1/2} \\
&\leqslant \delta \cdot g(t) + C_\delta\int_0^t \mathbb{E}|Z_\epsilon(s)|^{2p}\mathrm{d}s. \tag{14}
\end{aligned}
$$

Similarly, we have

$$\mathbb{E}\left|\int_0^t |Z_\epsilon(s)|^{2p} \cdot \|\dot{h}_\epsilon(s)\|_{l^2}\mathrm{d}s\right| \leqslant N\mathbb{E}\left(\int_0^t |Z_\epsilon(s)|^{4p}\mathrm{d}s\right)^{1/2}$$

6

$$\leqslant \quad \delta \cdot g(t) + C_{\delta,N} \int_0^t \mathbb{E}|Z_\epsilon(s)|^{2p} \mathrm{d}s. \qquad (15)$$

Letting $\delta = 1/4$ in (14) and (15) and combining the above calculations, we get

$$g(t) \leqslant g(0) + \frac{1}{2}g(t) + C\int_0^t \mathbb{E}|Z_\epsilon(s)|^{2p}\mathrm{d}s \leqslant 2g(0) + 2C\int_0^t g(s)\mathrm{d}s,$$

which gives the desired estimate by Gronwall's inequality. $\qquad \square$

Lemma 3.4. *For any $p \geqslant 1$ and $x \in \overline{D(A)}$, there exists $C_{p,T,N,x} > 0$ such that for any $\epsilon \in [0,1)$*

$$\mathbb{E}\left(\sup_{t\in[0,T]}|X^{\epsilon,h_\epsilon}(t,x)|^{2p}\right) + \mathbb{E}|K^{\epsilon,h_\epsilon}(\cdot,x)|_T^0 \leqslant C_{p,T,N,x}. \qquad (16)$$

Proof. First of all, as in the proof of Lemma 3.3 we can prove that

$$\mathbb{E}\left(\sup_{t\in[0,T]}|X^{\epsilon,h_\epsilon}(t,x)|^{2p}\right) \leqslant C_{p,T,N,x}. \qquad (17)$$

Let $a \in \mathrm{Int}(D(A))$ be as in Proposition 3.2. By Itô's formula, (11) and **(H2)**, we have

$$
\begin{aligned}
\frac{1}{2}|X^{\epsilon,h_\epsilon}(t) - a|^2 &= \frac{1}{2}|x - a|^2 + \int_0^t \langle X^{\epsilon,h_\epsilon}(s) - a, b(X^{\epsilon,h_\epsilon}(s))\rangle_{\mathbb{R}^m}\mathrm{d}s \\
&\quad + \int_0^t \langle X^{\epsilon,h_\epsilon}(s) - a, \sigma(X^{\epsilon,h_\epsilon}(s))\dot{h}_\epsilon(u)\rangle_{\mathbb{R}^m}\mathrm{d}s \\
&\quad + \sqrt{\epsilon}\int_0^t \langle X^{\epsilon,h_\epsilon}(s) - a, \sigma(X^{\epsilon,h_\epsilon}(s))\mathrm{d}W(s)\rangle_{\mathbb{R}^m} \\
&\quad - \int_0^t \langle X^{\epsilon,h_\epsilon}(s) - a, \mathrm{d}K^{\epsilon,h_\epsilon}(s)\rangle_{\mathbb{R}^m} \\
&\quad + \frac{\epsilon}{2}\int_0^t \|\sigma(X^{\epsilon,h_\epsilon}(s))\|_{L_2(l^2;\mathbb{R}^m)}^2\mathrm{d}s \\
&\leqslant \frac{1}{2}|x - a|^2 + C_b\int_0^t |X^{\epsilon,h_\epsilon}(s) - a|^2\mathrm{d}s \\
&\quad + \int_0^t \langle X^{\epsilon,h_\epsilon}(s) - a, b(a)\rangle_{\mathbb{R}^m}\mathrm{d}s \\
&\quad + N\left(\int_0^t |\sigma(X^{\epsilon,h_\epsilon}(s))^*(X^{\epsilon,h_\epsilon}(s) - a)|^2\mathrm{d}s\right)^{1/2} \\
&\quad + \sqrt{\epsilon}\int_0^t \langle X^{\epsilon,h_\epsilon}(s) - a, \sigma(X^{\epsilon,h_\epsilon}(s))\mathrm{d}W(u)\rangle_{\mathbb{R}^m} + \mu\gamma t \\
&\quad - \gamma|K^{\epsilon,h_\epsilon}|_t^0 + \mu\int_0^t |X^{\epsilon,h_\epsilon}(s) - a|\mathrm{d}s \\
&\quad + \frac{C_\sigma^2\epsilon}{2}\int_0^t (|X^{\epsilon,h_\epsilon}(s)| + \sigma(0))^2\mathrm{d}s.
\end{aligned}
$$

The desired estimate now follows by (17). $\qquad \square$

Define

$$w_\epsilon(t, x) := \int_0^t \sigma(X^h(s, x))(\dot{h}_\epsilon(s) - \dot{h}(s)) \mathrm{d}s. \tag{18}$$

The following lemma is easy by Ascoli-Arzela's lemma.

Lemma 3.5. $w_\epsilon(\cdot, x)$ converges a.s. to zero in $C([0, T], \overline{D(A)})$.

We now prove the following key lemma.

Lemma 3.6. X^{ϵ, h_ϵ} defined by (9) converges in probability to X^h defined by (10) in $\mathbb{S}$.

Proof. Set $v_\epsilon(t) := v_\epsilon(t, x) := X^{\epsilon, h_\epsilon}(t, x) - X^h(t, x)$. Then

$$
\begin{aligned}
v_\epsilon(t) &= K^{\epsilon, h_\epsilon}(t) - K^h(t) + \int_0^t (b(X^{\epsilon, h_\epsilon}(s)) - b(X^h(s))) \mathrm{d}s \\
&\quad + \int_0^t (\sigma(X^{\epsilon, h_\epsilon}(s)) \dot{h}_\epsilon(s) - \sigma(X^h(s)) \dot{h}(s)) \mathrm{d}s \\
&\quad + \sqrt{\epsilon} \int_0^t \sigma(X^{\epsilon, h_\epsilon}(s)) \mathrm{d}W(s).
\end{aligned}
$$

By Itô's formula, we have

$$
\begin{aligned}
|v_\epsilon(t)|^2 &= 2 \int_0^t \langle v_\epsilon(s), \mathrm{d}K^{\epsilon, h_\epsilon}(s) - \mathrm{d}K^h(s) \rangle_{\mathbb{R}^m} \\
&\quad + 2 \int_0^t \langle v_\epsilon(s), b(X^{\epsilon, h_\epsilon}(s)) - b(X^h(s)) \rangle_{\mathbb{R}^m} \mathrm{d}s \\
&\quad + 2 \int_0^t \langle v_\epsilon(s), (\sigma(X^{\epsilon, h_\epsilon}(s)) - \sigma(X^h(s))) \dot{h}_\epsilon(s) \rangle_{\mathbb{R}^m} \mathrm{d}s \\
&\quad + 2 \int_0^t \langle v_\epsilon(s), \sigma(X^h(s))(\dot{h}_\epsilon(s) - \dot{h}(s)) \rangle_{\mathbb{R}^m} \mathrm{d}s \\
&\quad + 2\sqrt{\epsilon} \int_0^t \langle v_\epsilon(s), \sigma(X^{\epsilon, h_\epsilon}(s)) \mathrm{d}W(s) \rangle_{\mathbb{R}^m} \\
&\quad + \epsilon \int_0^t \|\sigma(X^{\epsilon, h_\epsilon}(s))\|_{L_2(l^2; \mathbb{R}^m)}^2 \mathrm{d}s \\
&=: I_1^\epsilon(t) + I_2^\epsilon(t) + I_3^\epsilon(t) + I_4^\epsilon(t) + I_5^\epsilon(t) + I_6^\epsilon(t).
\end{aligned}
$$

It is clear that by (12)

$$I_1^\epsilon(t) \leqslant 0$$

and

$$I_2^\epsilon(t) \leqslant 2C_b \int_0^t |v_\epsilon(s)|^2 \mathrm{d}s.$$

By BDG's inequality and **(H2)** we also have

$$\mathbb{E}\left(\sup_{t \in [0,T]} |I_5^\epsilon(t)| \right) + \mathbb{E}\left(\sup_{t \in [0,T]} |I_6^\epsilon(t)| \right) \leqslant C\sqrt{\epsilon}.$$

As estimating (15) we have

$$\mathbb{E}\left(\sup_{s \in [0,t]} |I_3^\epsilon(s)| \right) \leqslant \frac{1}{2} \mathbb{E}\left(\sup_{s \in [0,t]} |v_\epsilon(s)|^2 \right) + C \int_0^t \mathbb{E}|v_\epsilon(s)|^2 \mathrm{d}s.$$

8

Set
$$g(t) := \mathbb{E}\left(\sup_{s\in[0,t]} |v_\epsilon(s)|^2\right).$$

Then we have
$$g(t) \leqslant \frac{1}{2}g(t) + C\sqrt{\epsilon} + \mathbb{E}\left(\sup_{s\in[0,t]} |I_4^\epsilon(s)|\right) + C\int_0^t \mathbb{E}|v_\epsilon(s)|^2 \mathrm{d}s,$$

which implies that
$$g(t) \leqslant C\sqrt{\epsilon} + 2\mathbb{E}\left(\sup_{s\in[0,T]} |I_4^\epsilon(s)|\right) + C\int_0^t g(s)\mathrm{d}s.$$

By Gronwall's inequality we get
$$\mathbb{E}\left(\sup_{s\in[0,T]} |v_\epsilon(s)|^2\right) \leqslant C\sqrt{\epsilon} + C\mathbb{E}\left(\sup_{s\in[0,T]} |I_4^\epsilon(s)|\right). \tag{19}$$

We now deal with the hard term I_4^ϵ. By Itô's formula again, we have
$$\begin{aligned}
\frac{1}{2}I_4^\epsilon(t) =&\ \langle v_\epsilon(t), w_\epsilon(t)\rangle_{\mathbb{R}^m} - \int_0^t w_\epsilon(s)\mathrm{d}(K^{\epsilon,h_\epsilon}(s) - K^h(s)) \\
&- \int_0^t w_\epsilon(s)(b(X^{\epsilon,h_\epsilon}(s)) - b(X^h(s)))\mathrm{d}s \\
&- \int_0^t w_\epsilon(s)(\sigma(X^{\epsilon,h_\epsilon}(s))\dot{h}_\epsilon(s) - \sigma(X^h(s))\dot{h}(s))\mathrm{d}s \\
&- \sqrt{\epsilon}\int_0^t w_\epsilon(s)\sigma(X^{\epsilon,h_\epsilon}(s))\mathrm{d}W(s) \\
=:&\ I_{41}^\epsilon(t) + I_{42}^\epsilon(t) + I_{43}^\epsilon(t) + I_{44}^\epsilon(t) + I_{45}^\epsilon(t).
\end{aligned}$$

For any $\delta > 0$ and $R > 0$, we have
$$\begin{aligned}
P\left(\sup_{t\in[0,T]} |I_{41}^\epsilon(t)| \geqslant \delta\right) =&\ P\left(\sup_{t\in[0,T]} |I_{41}^\epsilon(t)| \geqslant \delta; \sup_{t\in[0,T]} |v_\epsilon(t)| < R\right) \\
&+ P\left(\sup_{t\in[0,T]} |I_{41}^\epsilon(t)| \geqslant \delta; \sup_{t\in[0,T]} |v_\epsilon(t)| \geqslant R\right) \\
\leqslant&\ P\left(\sup_{t\in[0,T]} |w_\epsilon(t)| \geqslant \delta/R\right) + P\left(\sup_{t\in[0,T]} |v_\epsilon(t)| \geqslant R\right).
\end{aligned}$$

By Lemma 3.5 and (16) we know
$$\lim_{\epsilon\to 0} \mathbb{P}\left(\sup_{t\in[0,T]} |I_{41}^\epsilon(t)| \geqslant \delta\right) = 0.$$

Noting that
$$\sup_{t\in[0,T]} |I_{42}^\epsilon(t)| \leqslant \sup_{s\in[0,T]} |w_\epsilon(s)| \cdot (|K^{\epsilon,h_\epsilon}(\cdot)|_T^0 + |K^h(\cdot)|_T^0),$$

as above, we also have
$$\sup_{t\in[0,T]} |I_{42}^\epsilon(t)| \to 0 \text{ in probability.}$$

9

Similarly, we have

$$\sup_{t\in[0,T]} |I_{43}^\epsilon(t)| + \sup_{t\in[0,T]} |I_{44}^\epsilon(t)| \to 0 \text{ in probability.}$$

Moreover, by BDG's inequality we have

$$\sqrt{\epsilon}\mathbb{E}\left(\sup_{t\in[0,T]} |I_{45}^\epsilon(t)|\right) \leqslant C\sqrt{\epsilon}.$$

Combining the above calculations, we get

$$\sup_{t\in[0,T]} |I_4^\epsilon(t)| \to 0 \text{ in probability.}$$

It is easy to see by (16) that

$$\sup_{\epsilon\in(0,1)} \mathbb{E}\left(\sup_{t\in[0,T]} |I_4^\epsilon(t)|^2\right) < +\infty.$$

Hence

$$\lim_{\epsilon\to 0} \mathbb{E}\left(\sup_{t\in[0,T]} |I_4^\epsilon(t)|\right) = 0.$$

Substituting this into (19) we obtain

$$\lim_{\epsilon\to 0} \mathbb{E}\left(\sup_{s\in[0,T]} |v_\epsilon(s)|^2\right) = 0.$$

Thus, we have proven that for all $x \in \overline{D(A)}$

$$\sup_{t\in[0,T]} |v_\epsilon(t,x)|^2 \to 0, \quad \text{in probability.}$$

We now strengthen it by Lemma 3.3 to

$$\xi_{n,\epsilon} := \sup_{t\in[0,T],x\in D(A),|x|\leqslant n} |v_\epsilon(t,x)|^2 \to 0, \quad \text{in probability.}$$

Set

$$\mathbb{D}_n^\delta := \overline{D(A)} \cap \{x \in \mathbb{R}^m : |x| \leqslant n\} \cap \delta\mathbb{Z}^m,$$

where $\delta > 0$ and $\delta\mathbb{Z}^m$ denotes the grid in $\mathbb{R}^m$ with edge length δ. It is clear that there are only finite many points in $\mathbb{D}_n^\delta$. Hence

$$\xi_{n,\epsilon}^\delta := \sup_{t\in[0,T],x\in\mathbb{D}_n^\delta} |v_\epsilon(t,x)|^2 \to 0, \quad \text{in probability.}$$

For any $x \in \mathbb{R}^m$, let x_δ denote the left-lower corner point in $\delta\mathbb{Z}^m$ so that

$$|x - x_\delta| \leqslant \delta.$$

Noting that

$$\xi_{n,\epsilon} \leqslant 2\xi_{n,\epsilon}^\delta + 2 \sup_{x\in\overline{D(A)},|x|\leqslant n} \sup_{t\in[0,T]} |v_\epsilon(t,x) - v_\epsilon(t,x_\delta)|^2,$$

we have for any $\beta > 0$ and some $\alpha > 0$

$$\begin{aligned}
P(\xi_{n,\epsilon} > 4\beta) &\leqslant P(\xi_{n,\epsilon}^\delta + \sup_{x\in\overline{D(A)},|x|\leqslant n} \sup_{t\in[0,T]} |v_\epsilon(t,x) - v_\epsilon(t,x_\delta)|^2 \geqslant 2\beta) \\
&\leqslant P(\xi_{n,\epsilon}^\delta > \beta) + P\left(\sup_{x\in\overline{D(A)},|x|\leqslant n} \sup_{t\in[0,T]} |v_\epsilon(t,x) - v_\epsilon(t,x_\delta)|^2 > \beta\right)
\end{aligned}$$

10

$$\leqslant \quad P(\xi_{n,\epsilon}^\delta > \beta) + \mathbb{E}\left(\sup_{x \in \overline{D(A)}, |x| \leqslant n} \sup_{t \in [0,T]} |v_\epsilon(t,x) - v_\epsilon(t,x_\delta)|^2 \right) / \beta$$

$$\leqslant \quad P(\xi_{n,\epsilon}^\delta > \beta) + C\delta^\alpha / \beta,$$

where the last step is due to Lemma 3.3 and Kolmogorov's criterion. First letting δ be small enough, then ϵ go to zero, we then obtain

$$\lim_{\epsilon \to 0} P(\xi_{n,\epsilon} > 4\beta) = 0.$$

which yields the desired result. $\qquad \square$

Lemma 3.7. (LD)$_1$ *holds.*

Proof. Let h_ϵ be a sequence in $\mathcal{A}_N^T$ converge to h in distribution. Since $\mathcal{D}_N$ is compact and the law of W is tight, $\{h_\epsilon, W\}$ is tight in $\mathcal{D}_N \times \mathbb{C}_T(\mathbb{U})$ by the definition of tightness. Without loss of generality, we assume the law of $\{h_\epsilon, W\}$ weakly converges to μ. Then the law of h is just $\mu(\cdot, \mathbb{C}_T(\mathbb{U}))$. Indeed, for any bounded continuous function g on $\mathcal{D}_N$, we have

$$\mathbb{E}(g(h)) = \lim_{n \to \infty} \mathbb{E}(g(h_\epsilon)) = \int_{\mathcal{D}_N} g(h)\mu(\mathrm{d}h, \mathbb{C}_T(\mathbb{U})).$$

By Skorohod's representation theorem, there are $(\tilde{\Omega}, \tilde{P})$ and $\{\tilde{h}_\epsilon, \tilde{W}^\epsilon\}$ and $\{\tilde{h}, \tilde{W}\}$ such that
  (1) $(\tilde{h}_\epsilon, \tilde{W}^\epsilon)$ a.s. converges to $(\tilde{h}, \tilde{W})$;
  (2) $(\tilde{h}_\epsilon, \tilde{W}^\epsilon)$ has the same law as (h_ϵ, W);
  (3) The law of $\{\tilde{h}, \tilde{W}\}$ is μ, and the law of h is the same as $\tilde{h}$.
  Using Lemma 3.6, we get

$$\Phi(\frac{1}{\sqrt{\epsilon}} \int_0^{\cdot} \dot{\tilde{h}}_\epsilon \mathrm{d}s + \tilde{W}^\epsilon) \to X^{\tilde{h}}, \quad \textit{in probability},$$

where Φ is the strong solution functional(cf. [9]). From this, we derive

$$\Phi(\frac{1}{\sqrt{\epsilon}} \int_0^{\cdot} \dot{h}_\epsilon \mathrm{d}s + W) \to X^h, \quad \textit{in distribution}.$$

Thus, **(LD)$_1$** holds. $\qquad \square$

Similar to the proof of Lemma 3.6, one can easily verify that

Lemma 3.8. (LD)$_2$ *holds.*

Thus, by Lemmas 3.7, 3.8 and Theorem 2.5, we have proved Theorem 3.1.

REFERENCES

[1] Anderson, R.F. and Orey, S.: Small Random Perturbation of Dynamical Systems with Refecting Boundary. Nagoya Math. J., Vol.60, 189-216(1976).
[2] Boué, M. and Dupuis, P.: A variational representation for certain functionals of Brownian motion. *Ann. of Prob.*, Vol. 26, No.4, 1641-1659(1998).
[3] Budhiraja, A. and Dupuis, P.: A variational representation for positive functionals of infinite dimensional Brownian motion. *Probab. Math. Statist.* 20, no.1, Acta Univ. Wratislav. No. 2246, 39–61(2000).
[4] Budhiraja, A., Dupuis, P. and Maroulas, V.: Large deviations for infinite dimensional stochastic dynamical systems, to appear in *Ann. of Prob.*.

[5] Cépa, E.: Équations différentielles stochasticcques multivoques. Lect. Notes in Math. *Sém. Prob. XXIX*, Springer, Berlin, 86-107(1995).

[6] Cépa, E.: Probleme De Skorohod Multivoque. Ann. of Prob., Vol.26, No.2, 500-532(1998).

[7] Deuschel, J.D. and Stroock, D.W.: Large deviations. Academic Press, Boston, New York, 1988.

[8] Dupuis, P. and Ellis, R.S.: *A Weak Convergence Approach to the Theory of Large Deviations.* Wiley, New-York, 1997.

[9] Ikeda, N. and Watanabe S.: *Stochastic Differential Equations and Diffusion Processes*, 2nd ed., Kodansha, Tokyo/North-Holland, Amsterdam, 1989.

[10] Liu, W.: Large deviations for stochastic evolution equations with small multiplcative noise. Preprint.

[11] Ren, J. and Zhang, X.: Freidlin-Wentzell's large deviations for homeomorphism flows of non-Lipschitz SDEs. *Bull. Sci. Math. 2 Serie*, Vol 129/8 pp 643-655(2005).

[12] Ren, J. and Zhang, X.: Freidlin-Wentzell's Large Deviations for Stochastic Evolution Equations. *J. Func. Anal.*, Vol.254, 3148-3172(2008).

[13] Xu, S.: Explicit solutions for multivalued stochastic differential equations, Statist. Probab. Lett. 78, 2281-2292(2008).

[14] Zhang, X.: Skorohod problem and multivalued stochastic evolution equations in Banach spaces. Bull. Sci. Math.131(2),175-217(2007).

[15] Zhang, X.: Euler Schemes and Large Deviations for Stochastic Volterra Equations with Singular Kernels. *Journal of Differential Equation*, Vol. 224, 2226-2250(2008).

[16] Zhang, X.: Stochastic Volterra Equations in Banach Spaces and Stochastic Partial Differential Equations. Preprint.

[17] Zhang, X.: A variational representation for random functionals on abstract Wiener spaces. Preprint.